TDB-ACC-NO: NB9209418

DISCLOSURE TITLE: Estimating the Acoustic Scores in a Speech Recognition System Using Context Dependent Acoustic Markov Models..

PUBLICATION-DATA: IBM Technical Disclosure Bulletin, September 1992, US

VOLUME NUMBER: 35

ISSUE NUMBER: 4B

PAGE NUMBER: 418

PUBLICATION-DATE: September 1, 1992 (19920901)

CROSS REFERENCE: 0018-8689-35-4B-418

DISCLOSURE TEXT:

- In one prominent approach to speech recognition, context-dependent acoustic Markov models are used. Corresponding to each arc in the inventory of arcs is a collection of context-dependent prototypes: the appropriate prototype is selected according to the context of the arc. The mapping from context to prototype is not always optimal - especially if it is obtained empirically from training data. Inaccuracies in the mapping and/or the prototypes can lead to a poor score being ascribed to a frame of speech, when a good score would be obtained if a different prototype of the same arc were used instead. The invention below takes advantage of this observation to smooth the acoustic scores, thereby reducing the recognition error rate.

- The acoustic score associated with arc A in context C for vector X is computed as follows.

(1) Using the context-to-prototype mapping, determine the prototype P which represents arc A in context C.

(2) Compute the acoustic score of vector X when modelled by the prototype P of Step 1. Denote this score as S.

(3) Perform Step 4 for each prototype Q associated with arc A.

(4) Compute the acoustic score of vector X when modelled by prototype Q.

(5) Locate the maximum score obtained in Step 5 and denote it as M.

(6) Compute a smoothed score for vector X and arc A in context C as W.S + (1 - W).M

Reasonable values for W lie in the range 0.9 to 0.99.

VARIATIONS

Steps 5 and 6 above may be replaced by the following steps instead.

(7) Sum the scores obtained in Step 5 and denote the total as T.

(8) Compute a smoothed score for vector X and arc A in context C as

Z.
S + (1 - Z).T
Reasonable values for Z lie in the range 0.9 - 0.999.